A Gravity Model of International Student Flows

Luciana Mendoza Vargas & Samuel Schreyer, Ph.D., Fort Hays State University

Abstract

Our research aims to use a gravity model to assess the importance of geographical distance and other macroeconomic factors associated with the destination country chosen by students when pursuing tertiary education abroad. We analyze a large sample of data that includes international students from nearly every country in the world from 2013 to 2021. Our results indicate that geographical distance is an important factor when students choose to study abroad and that this effect is most pronounced for students from lower-income countries. We also find that the flow of international students became more sensitive to geographical distance as the years progressed in our sample. We speculate this is occurring because a growing proportion of international students are from lower-income countries and are more sensitive to costs incurred with longer geographical distances.

Introduction

International student flows have grown tremendously over the past several decades. In 2022, about six million international students were enrolled in universities, nearly four times the number in 1999 (OECD, 2023). Significant macroeconomic benefits are associated with increased international student flows. For destination countries, international students who continue to reside and work in the country after graduation are a source of skilled labor and can also offset declines in their domestic labor force. The countries of origin, too, benefit by having these workers send remittances to family and friends in their home country. Similarly, the countries of origin see increased human capital and technological know-how when students return home and work.

Methodology & Model

Jan Tinbergen (1962) described bilateral trade flows as being similar to Isaac Newton's universal law of gravitation in that the volume of trade between two countries is proportional the size of their economies and inversely proportional to the distance between them. Now called the gravity model, this modelling approach has become ubiquitous in the field of international trade. Economists have also used the gravity model to describe labor migration patterns between countries. Our study follows several recent studies who adapt the gravity model to examine international student flows (e.g., Abbott and Silles, 2016).

Our research is distinguished from previous studies in two primary ways. First, we use a vast dataset covering 205 origin countries and 42 destination countries from 2013 to 2021. Our data is retrieved from the World Bank, Organization for Economic Co-operation and Development, Times Higher Education, and the Centre d'Etudes Prospectives et d'Informations Internationales. Second, our model uses interaction terms to explore how the relationship between geographical distance and student flows has changed over time.

Selected Results

Figure 1 shows the number of students pursuing tertiary education by country of origin. The top three countries are China, India, and Germany, which account for 35% of all international students in our sample. **Figure 2** lists the primary covariates we investigate in our study and the regression results. Regression (1) uses the full sample of data, whereas regression (2) limits the sample to origin countries classified as low-income or low-middle income. The estimated coefficient for *distance* in these regressions suggests that a 10% increase in distance between countries is associated with a 10-16% decrease in student flows. The results also indicate that a common language, religion, and colonial past are important factors governing student flows.

Regression (3) builds on the previous regression by adding interaction terms for *distance* and the year dummies. The interactions terms are statistically significant, but not shown for brevity. **Figure 3** shows predicted students flows from regression (3) for a hypothetical pair of origin and destination countries each year from 2013 to 2021. Note that although the volume of international students is predicted to increase each year, the relationship with *distance* is increasingly negatively sloped meaning that student flows have become more sensitive to geographical distance. We find this point noteworthy because it runs counter to the notion that technological advances are reducing the disadvantages of geographical distance for international students. We speculate our finding occurs because a growing share of international students are coming from low-income and low-middle income countries and are more sensitive to costs associated with geographical distance.

Conclusion

Understanding international student flows is important because they are associated with significant macroeconomic effects for the countries involved. This study uses a gravity model to estimate the effect that geographical distance and other factors have on the flow of students. We find that two countries with a common language, common religion, a shared colonial past, and a shorter distance between tend to have more international students. We do not find evidence that geographical distance is becoming less important in explaining student flows. Indeed, our findings show that student flows originating from low and low-middle-income countries have become more sensitive to distance in recent years.

Selected References

Abbott, A., & Silles, M. (2016). Determinants of international student migration. *The World Economy*, *39*(5), 621-635.

OECD (2023), *International Migration Outlook 2023*, OECD Publishing, Paris.

Tinbergen, J. (1962). Shaping the world economy; suggestions for an international economic policy.

Figure 1. Number of Students Studying Abroad (Yearly Average, 2013-2021)

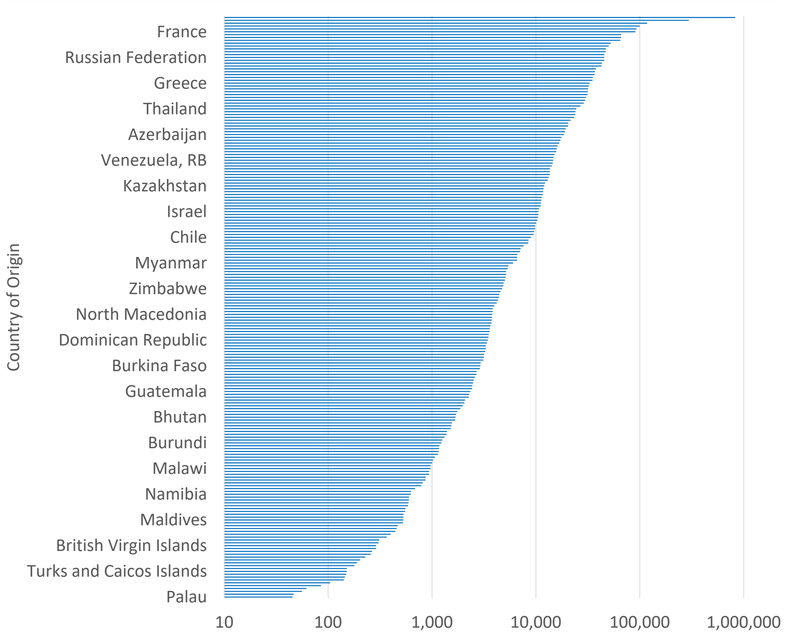

Figure 2. Fixed Effects Regressions

Variable	Description	(1)	(2)	(3)
$distance_{od}$	Distance (kilometers) between the most populated cities in o and d. Log transformed.	-1.044*** (0.0382)	-1.571*** (0.0817)	-1.498*** (0.0928)
GDP_d/GDP_o	GDP per capita in country d divided by that in o, in 2015 USD. Lagged one year.	-0.00501*** (0.00135)	-0.000265 (0.00158)	-0.000272 (0.00158)
$u_d - u_o$	Rate of unemployment in country d less that in o. Lagged one year.	-0.0134*** (0.00271)	-0.0119** (0.00510)	-0.00977* (0.00508)
$score_d$	Mean reputational score of universities in country d. Log transformed and lagged one year.	-0.331*** (0.0361)	-0.548*** (0.0607)	-0.558*** (0.0606)
$border_{od}$	Dummy equal to 1 if countries d and o share a land border, otherwise 0.	-0.0629 (0.173)	0.117 (0.814)	0.103 (0.812)
$language_{od}$	Dummy equal to 1 if countries d and o share a common language, otherwise 0.	0.973*** (0.0750)	1.097*** (0.106)	1.096*** (0.106)
$religion_{od}$	Religious proximity index calculated as product of share of Catholics, Protestants, and Muslims in countries d and o. The index ranges between 0 and 1.	1.230*** (0.112)	1.354*** (0.200)	1.355*** (0.200)
$colony_{od}$	Dummy equal to 1 if countries d and o have had a colonial relationship since WWII, otherwise 0.	1.795*** (0.188)	2.018*** (0.225)	2.019*** (0.225)
Interaction terms	Defined as the product of $distance_{od}$ with each year dummy.	No	No	Yes
Observations		37,991	14,851	14,851
R-squared		0.731	0.734	0.734

Note: Letters o and d refer to the origin and destination countries, respectively. The dependent variable is the number of students from o pursuing tertiary education in d, measured in logs. Regression (1) uses the full sample, while regressions (2) and (3) limit the sample to countries of origin classified by the World Bank as low income or low-middle income. For brevity, estimates for the fixed country and time effects are now shown. Robust standard errors in parentheses. Statistical significance indicated as: *** p<0.01, ** p<0.05, * p<0.1.

Figure 3. The Effect of Geographical Distance on Student Flows By Year